ARCON Corporation
J.D. Tardelli - jdtardelli@arcon.com

# ASE Test & Evaluation Phase 1 & 2


Distribution Statement A: Approved for Public Release, Distribution Unlimited

ARCON Corporation
J.D. Tardelli - jdtardelli@arcon.com

# Overview

- **Phase 1 Pilot Corpus Review**
- **Phase 1 Test & Evaluation**

- **Phase 2 Demonstration Corpus**
- **Proposed Phase 2 Testing**
- **Proposed Phase 2 Corpus Plans**
  - ***Pilot Corpus***
  - ***Study Corpus***
  - ***Test Corpus***
  - ***Proponent Corpus Collection***

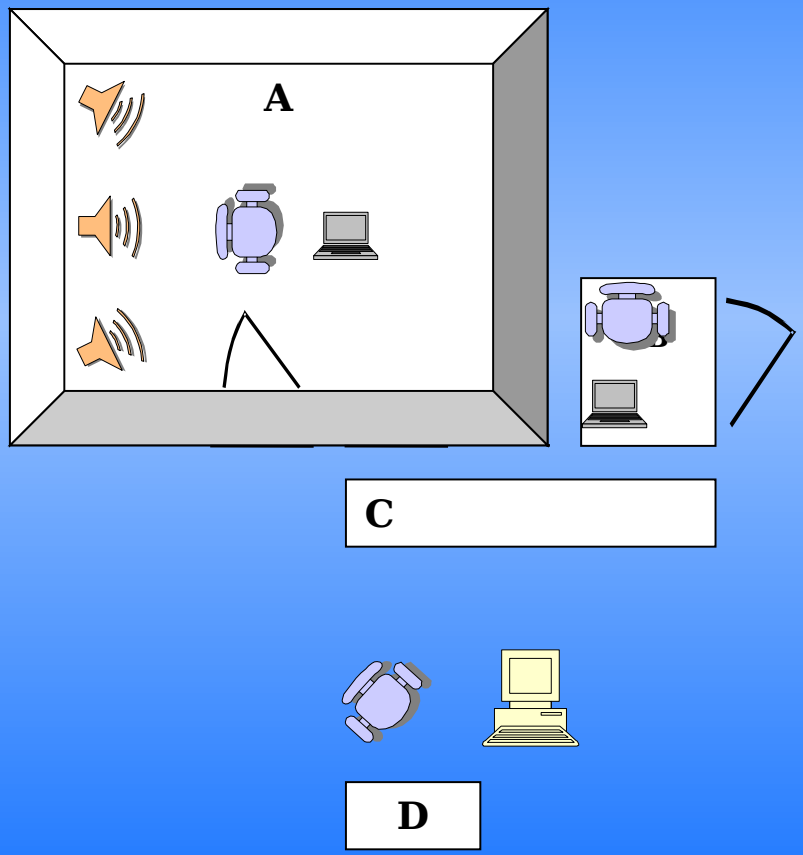
ARCON Corporation
J.D. Tardelli - jdtardelli@arcon.com

# Phase 1 Pilot Corpus

- ***Research and Test Components***
- ***10 Male + 10 Female Talkers***
- ***Vowel, Word, Sentence and Conversational Material***
- ***9 Acoustic Noise Environments***
- ***7/8 Sensor Channels***
  - ***2 GEMS channels***
  - ***1 EGG***
  - ***2 P-mic sensors (throat and fore head)***
  - ***1/2 Resident microphone***
  - ***1 Calibration microphone***
- ***70 Gbytes***

ARCON Corporation
J.D. Tardelli - jdtardelli@arcon.com

# Facility Description




**A: Recording Chamber (12’ x 10’)**
- **Stereo Speaker + Subwoofer**
- **PC List Presentation**
- **Sensors**
- **Intercom**
- **Patch Panel**

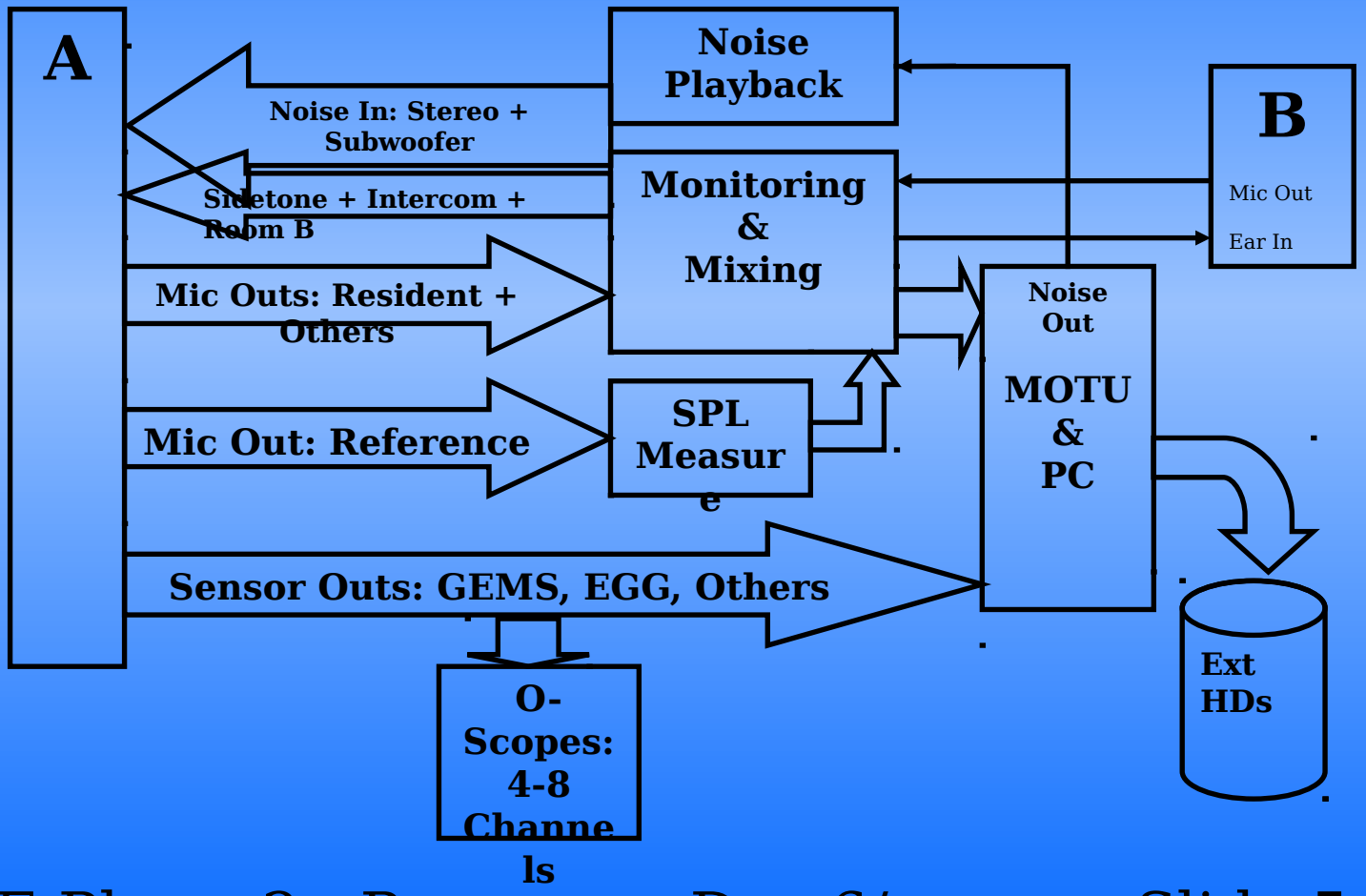
**B: Chamber 2 Conversational (2’ x 3’)**
- **Speaker Array**
- **Intercom**
- **Patch Panel**

**C: Sound Field Measurement**

**D: Audio Engineering Station**
- **PC Workstation, MOTU and HDs**
- **Sound Field Reconstruction**
- **Signal Generators and**

ARCON Corporation
J.D. Tardelli - jdtardelli@arcon.com

# Playback & Recording



ARCON Corporation
J.D. Tardelli - jdtardelli@arcon.com

# *Multitrack Synchronization*

ARCON Corporation
J.D. Tardelli - jdtardelli@arcon.com

# *Pilot Corpus Problems*

- Sensor Signal Consistency
- Sensor Placement Limitations
- Sensor / Talker Interactions
- Talker Variability / Pathologies
- Limited Acoustic Field Frequency Range

ARCON Corporation
J.D. Tardelli - jdtardelli@arcon.com

# Phase 1 Test & Evaluation

- ***Characteristics - Methodologies***
  - Intelligibility - DRT
  - Quality - Pair Comparison (A/B)
- ***Baseline Evaluation***
  - ***Intelligibility***
- ***Final Evaluation***
  - ***Intelligibility***
  - ***Quality***

ARCON Corporation
J.D. Tardelli - jdtardelli@arcon.com

# *Phase 1 Baseline Test & Evaluation*

- **Intelligibility - DRT**
  - **One DRT Lists for Each Talker in 9 Environments**
  - **Null and MELPe Processed**
  - **Select Talkers for Evaluation DRTs**
  - **Analysis of Baseline DRT Results**
    - Null Analysis
    - MELPe Analysis

ARCON Corporation
J.D. Iardelli - jdiardelli@arcon.com

# ASE Phase 1 Final Test & Evaluation

- **Intelligibility - DRT**
  - Number Participants - 3
  - Environments M2H, BHH, MOH
    - All participants not test in all environments
  - Standard DRT Practices with no less than 10 Listeners, results reported for “culled” set of 8 Listeners.
  - Talkers 3M/3F from Pilot or Supplemental Corpora
    - Talker selection limited for some participants
  - 40dB Criteria - $I_{Participant}$ (high) $\geq$ $I_{MELPe}$ (low)

# Phase 1 Final Test & Evaluation (cont.)

ARCON Corporation
J.D. Tardelli - jdtardelli@arcon.com

- **Quality - Forced Choice Paired Comparison A/B Test**
  - Number Participant Systems - 3
    - with CELP, MELPe(2400), MELPe(1200) as controls
  - Environments M2H, BHH, MOH + BHL
    - All participants not test in all environments
  - Talkers 3M/3F from Pilot or Supplemental Corpora
    - Talker selection limited for some participants
  - 48 Subjects as 6 Groups of 8 Subjects
  - # Stimuli = 6 talkers * 4 env. * 6 coders * 2 orders * 2 presentations = 576
  - 96 minutes testing/group, 6 16-minute blocks of 96 stimuli each
  - Standard Test Procedures re randomization and block balance
  - Criteria - $Q_{Participant}(high) \geq Q_{CELP}(high)$

ARCON Corporation
J.D. Iardelli - jdtardelli@arcon.com

# Phase 2 Demonstration Corpus

- Subset of Pilot Corpus
- 2 Male + 2 Female Talkers from “test” set
- DRT, Harvard PB Sentence & Vowel Lists
- Blackhawk, Bradley M2 and MOUT
  - High and Low Noise States
- All Sensor Channels
- Supplemental Material
- Provided Under Limited License to aid in BAA response

# Phase 2 Demonstration Corpus (cont.)

ARCON Corporation
J.D. Tardelli - jdtardelli@arcon.com

- Supplemental Materials
  - Acoustic noise files for the six environments presented (MBO)
  - All talker demographic and sensor information
  - Lists of all scripted material used in this data base
  - DRT word list decode information
  - Microphone and sensor information
  - ARCON Pilot Corpus Phase 1&2 design review presentations
  - ARCON Pilot Corpus Final Report
  - A short “Cheat Sheet” with environment and sensor codes
  - A list of sensor file problems and exceptions
  - DRT Baseline results
    - All Talkers
    - NULL and MELPe 2400 processed
  - Analysis of Baseline DRT results

ARCON Corporation
J.D. Tardelli - jdtardelli@arcon.com

# Proposed Phase 2 Testing

- Program Milestone and Waypoint Testing
  - Supported by Test Corpus
  - System Characteristics to be Tested
    - Intelligibility by DRT
    - Quality by Paired Comparison (A/B)
    - Aural Speaker Recognition, “Recognizability” by Schmidt-Nielsen Method
- Proponent Testing
  - Encouraged

ARCON Corporation
J.P. Tardelli - jptardelli@arcon.com

# Aural Speaker Recognition

- Recognition of the Source Talker by the Receive Listener
- Measured at the output of the synthesizer
- What Speech?
  - Coded Speech vs. Coded Speech
  - Coded Speech vs. Direct Speech
- Subjective SAME - DIFFERENT Method
- Using Scripted Materials

ARCON Corporation
J.D. Tardelli - jdtardelli@arcon.com

# Proposed Phase 2 Corpus Plans

- ***Demonstration Corpus***
  - limited license for use associated with BAA response
- ***Phase 1 Pilot Corpus***
  - to be released to Phase 2 Participants
- ***Study Corpus***
  - Talker Variability; June/July Collection
- ***Test Corpus***
- ***Proponent Corpus Collection***

ARCON Corporation
J.D. Tardelli - jdtardelli@arcon.com

# *Study Corpus*

- Talker Variability
  - regional dialects
  - recognizability characteristics
- Sensor Electronics
- Optimized Sensor Placement
- Expanded Frequency Range of Acoustic Noise Simulations
- Aid in Design of ASE P2 Test Corpus

ARCON Corporation
J.D. Jensen jdjensen@arcon.com

# Proposed Phase 2 Test Corpus

- **Design Criteria**
  - This is a Test Corpus not a Research Corpus.
  - The Test Corpus needs to support all of the test methodologies of ASE P2.
  - ASE P2 Participant's proprietary sensors need to be considered from the start.
  - With either Waypoint or Milestone testing every 6 months, the Test Corpus will need to be supplemented as the test program proceeds through its stages..
  - Talker effects need to be controlled.
  - At certain stages of the program the Test Corpus will need to be segmented into releasable Training and private Test data sets.
  - Parts of the Test Corpus must be available in a timely manor in order to meet both participant requirements and the test schedule.
  - The collection of the Test Corpus must be cost and time effective.

ARCON Corporation
I.D. Tardelli - idtardelli@arcon.com

# *Proponent Corpus Collection*

- **Encouraged**
- **Speech + Noise**
- **Human Subject Issues**
- **Sensor Factors**
- **Synchronous Recording**

ARCON Corporation
J.D. Tardelli - jdtardelli@arcon.com

# Speech + Acoustic Noise

- Mixing Methods
  - Electronic
  - Head & Torso Simulator in Noise Simulation
  - Subject Recording in Noise Simulation
  - In Situ Recording
- Acoustic Noise Field Sources
  - Demonstration Corpus
  - ARCON

ARCON Corporation
J.D. Tardelli - jdtardelli@arcon.com

# Human Subject Issues

- Risk Factors
  - Active Sensors, Acoustic Noise Exposure
- Minimization of Risk Factors
  - Subject Instrumentation
    - # Sensors, Placement, Noise Attenuation
  - Subject Recording Protocols
    - Hearing Thresholds
    - Cumulative Noise and Active Sensor Exposure
    - Schedule
- Subject Data Base
  - Demographics +
  - Privacy

ARCON Corporation
I.D. Tardelli – idtardelli@arcon.com

# Human Subject Issues (cont.)

- Informed Consent Document
  - Procedures to safeguard personal information
  - Details on minimization of risk
    - Inputs from Sensor Manufacturers
    - OSHA Noise Exposure Limits
- Independent Review Board
  - Initial review of all documentation,
  - Possible modifications,
  - Annual review and
  - Intermediate reviews of any changes to protocols or practices.

ARCON Corporation
J.D. Tardelli - jdtardelli@arcon.com

# Record Session Noise Exposure

- Sample Scripted Material Schedule
  - Level Adjustment Materials (.5 min)
  - Sustained Vowels within Words (1 min)
  - Diagnostic Rhyme Test Word Lists (8 min)
  - Harvard Sentence Lists (3 min)
  - CVC Nonsense Words in Carrier Phrase (3 min)
- Noise Exposure Schedule

ARCON Corporation
J.D. Tardelli - jdtardelli@arcon.com

| Subject Noise Exposure Schedule | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Condition | Time (min) | SPL dBA | OSHA Max (hrs) | Unattenuated Daily Dose | Attenuation dBA | Effective Noise Level dBA | OSHA Max (hrs) | Attenuated Daily Dose |
| Office Environment | | | | | | | | |
| Exposure | 20 | 45.0 | 4096.0 | 0.01% | 0.0 | 45.0 | 4096.0 | 0.01% |
| Break | 30 | | | | | | | |
| MCE Environment | | | | | | | | |
| Exposure | 20 | 66.0 | 222.9 | 0.15% | 34.4 | 31.6 | 26249.5 | 0.00% |
| Break | 35 | | | | | | | |
| M2 Environment (High) | | | | | | | | |
| Exposure | 20 | 104.0 | 1.1 | 29.02% | 25.9 | 78.1 | 41.6 | 0.80% |
| Break | 30 | | | | | | | |
| M2 Environment (Low) | | | | | | | | |
| Exposure | 20 | 64.0 | 294.1 | 0.11% | 25.9 | 38.1 | 10660.6 | 0.00% |
| Long Break | 65 | | | | | | | |
| MOUT Environment (High) | | | | | | | | |
| Exposure | 20 | 98.0 | 2.6 | 12.63% | 7.0 | 91.0 | 7.0 | 4.79% |
| Break | 30 | | | | | | | |
| MOUT Environment (Low) | | | | | | | | |
| Exposure | 20 | 58.0 | 675.6 | 0.05% | 7.0 | 51.0 | 1782.9 | 0.02% |
| Break | 35 | | | | | | | |
| UH-60 Environment (High) | | | | | | | | |
| Exposure | 20 | 98.0 | 2.6 | 12.63% | 20.0 | 78.0 | 42.2 | 0.79% |
| Break | 30 | | | | | | | |
| UH-60 Environment (Low) | | | | | | | | |
| Exposure | 20 | 65.0 | 256.0 | 0.13% | 20.0 | 45 | 4096.0 | 0.01% |
| Daily Cumulative Noise Exposure | | | | 54.73% | | | | 6.42% |

ARCON Corporation
J.D. Tardelli jdtardelli@arcon.com

# Sensor Factors

- **Calibration Microphones**
- **Resident Microphones**
- **GEMS Interference**
- **P-Mic Pre-amp**

# Synchronous Recording

- **Use of multiple MOTUs**

ARCON Corporation
J.D. Tardelli - jdtardelli@arcon.com

# Supporting Material

ARCON Corporation
J.B. Tardelli - jtardelli@arcon.com

# Scripted Material - Sustained Vowels

| sound | word |
|---|---|
| ^ | h<u>u</u>t |
| a: | h<u>a</u>rm |
| @ | h<u>a</u>t |
| e | h<u>e</u>ck |
| e: | h<u>u</u>rt |
| i | h<u>i</u>t |
| i: | h<u>ea</u>t |
| o | h<u>o</u>t |
| o: | h<u>o</u>re |
| u | h<u>oo</u>d |
| ai | h<u>i</u>ve |
| au | h<u>ow</u> |
| Ou | h<u>o</u>me |
| e.. | h<u>ai</u>r |
| ei | h<u>a</u>y |
| i.. | h<u>e</u>re |
| oi | h<u>oy</u> |

ARCON Corporation
J.D. Tardelli - jdtardelli@arcon.com

# Scripted Material - Diagnostic Rhyme Test (DRT) Word Lists

| | | | |
|---|---|---|---|
| **MOOT** | **or** | **BOOT** | **Voicing** |
| **SHEET** | **or** | **CHEAT** | **Nasality** |
| **JAB** | **or** | **GAB** | **Sustention** |
| **POT** | **or** | **TOT** | **Sibilation** |
| **GHOST** | **or** | **BOAST** | **Graveness** |
| **DINT** | **or** | **TINT** | **Compactness** |

ARCON Corporation
J.D. Tardelli - jdtardelli@arcon.com

# Scripted Material - Harvard Sentences

## 1965 Harvard Phonetically Balanced Sentences

## 72 lists of 10 sentences balanced across lists

**List 1:**

1. THE BIRCH CANOE SLID ON THE SMOOTH PLANKS.
2. GLUE THE SHEET TO THE DARK BLUE BACKGROUND.
3. IT'S EASY TO TELL THE DEPTH OF A WELL.
4. THESE DAYS A CHICKEN LEG IS A RARE DISH.
5. RICE IS OFTEN SERVED IN ROUND BOWLS.
6. THE JUICE OF LEMONS MAKES FINE PUNCH.
7. THE BOX WAS THROWN BESIDE THE PARKED TRUCK.
8. THE HOGS WERE FED CHOPPED CORN AND GARBAGE.
9. FOUR HOURS OF STEADY WORK FACED US.
10. A LARGE SIZE IN STOCKINGS IS HARD TO SELL.

ARCON Corporation
J.D. Tardelli - jdtardelli@arcon.com

# Conversational Scenarios

- Designed to:
  - be asymmetric with the talker providing the majority of the verbal interaction
  - last a minimum of five minutes
  - minimally non-combative and non-controversial
  - of general interest regardless of experience, sex, or age
  - designed to reduce fatigue and increase interest
  - be quickly read and learned and not require special skills to implement

ARCON Corporation
I.D. Tardelli - idtardelli@arcon.com

# *Acoustic Noise Environments*

- **Blackhawk Helicopter (98dBA, 112dBC SPL)**
- **M2 Bradley APC (109-112dBA SPL)**
- **Military Operations in Urban Terrain - MOUT (40dBA floor, 98dBA gunfire)**

ARCON Corporation
J.D. Tardelli - jdtardelli@arcon.com

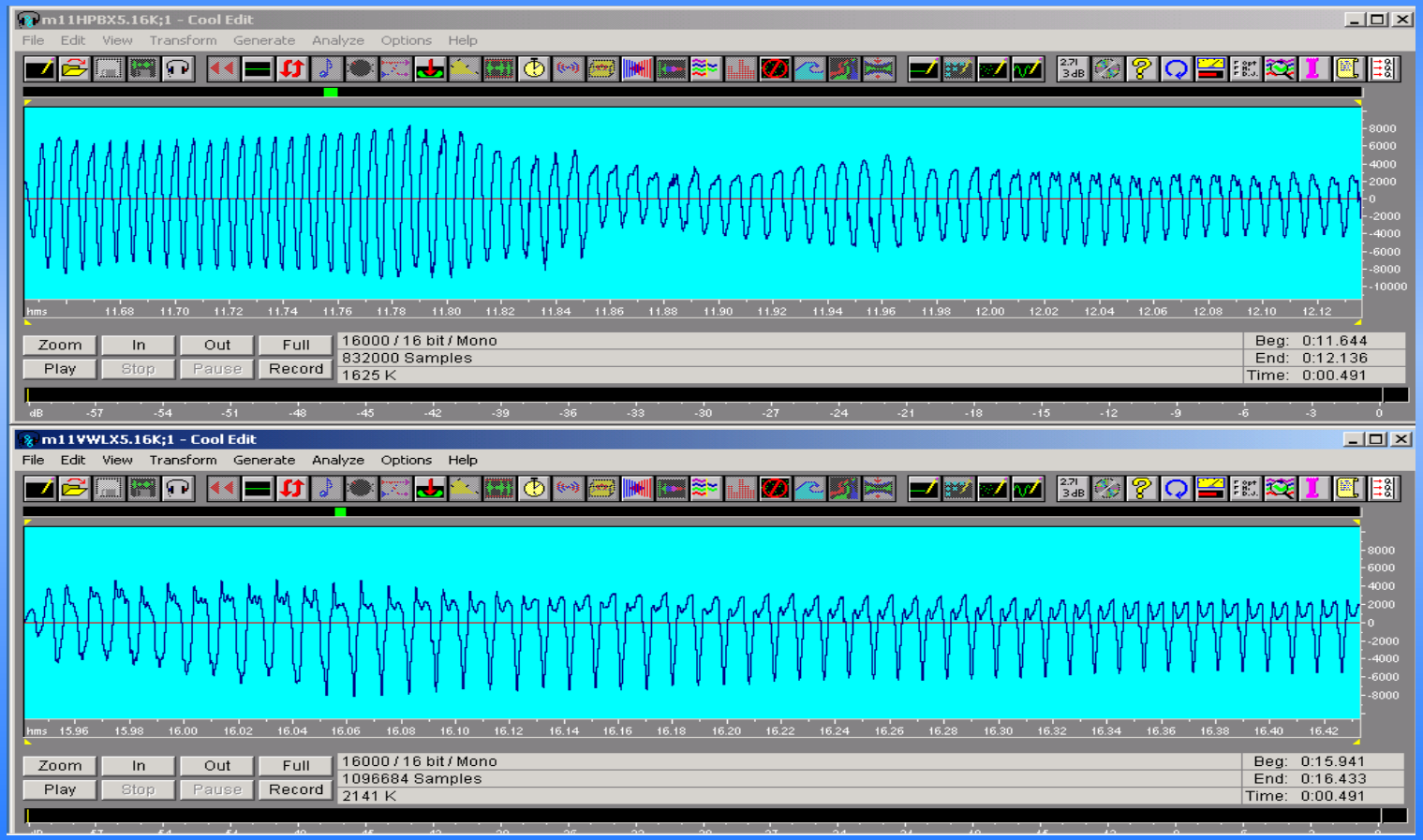
ARCON Corporation
J.D. Tardelli - jdtardelli@arcon.com

ARCON Corporation

# Sensor Attachment & Placement

ARCON Corporation
J.D. Tardelli, tardelli@arcon.com

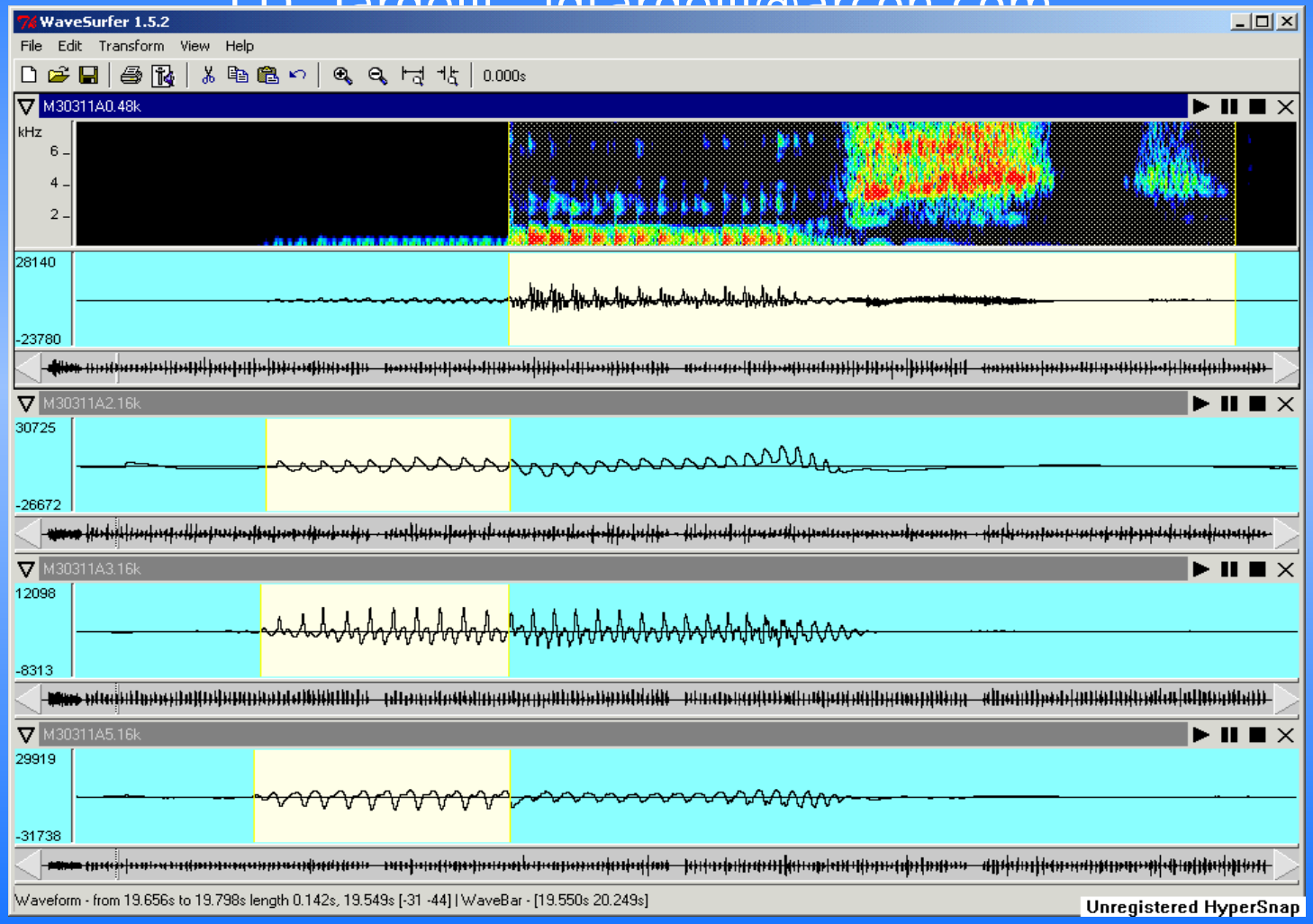
# GEMS Signal Example




Harvard Sentence vs. Vowel

ARCON Corporation
J.D. Tardelli - jdtardelli@arcon.com

# *Voice Bar Example - M3 Quiet - "BOAST"*




Resident Mic

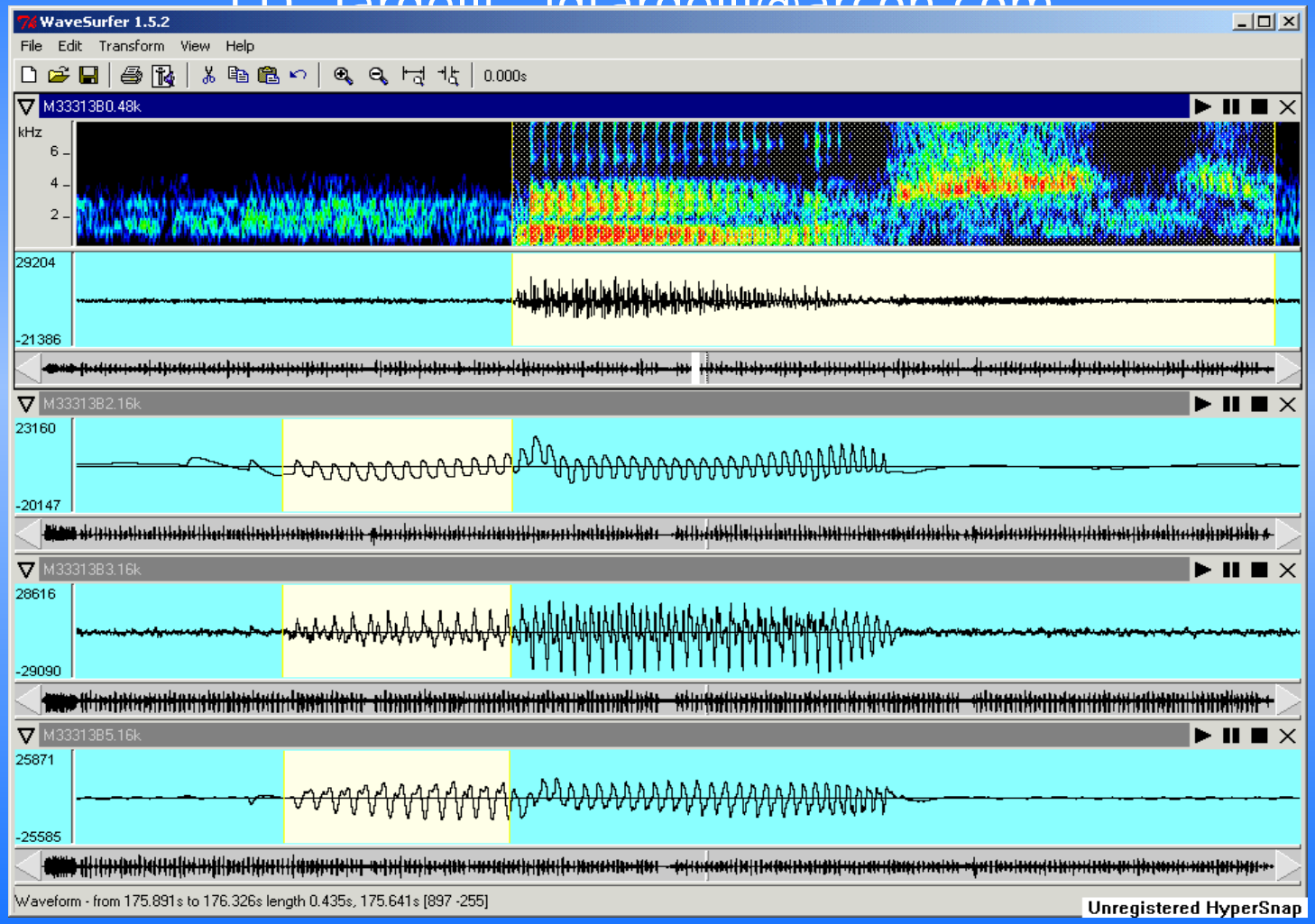
EGG

P-Mic Throat

GEMS

ARCON Corporation
J.D. Tardelli - jdtardelli@arcon.com

# *Voice Bar Example - M3 M2H - "BOAST"*




Resident Mic

EGG

P-Mic Throat

GEMS

ARCON Corporation
J.D. Tardelli - jdtardelli@arcon.com

# Resident Transducers

- **M2 Bradley**
  - **CVC Helmet with Active Noise Reduction (ANR) Bose Product Improved Combat Vehicle Crewman (PICVC) Helmet, H-374(V)5/VRC; Second generation ANR helmet with M175 Electret noise canceling boom microphone assembly.**
- **MOUT**
  - **MICH Helmet with Communications Subsystem: The communication subsystem is intended to provide aural protection, situational awareness and a dual channel communications capability.**

ARCON Corporation
J.D. Tardelli - jdtardelli@arcon.com

SYSTEM TESTED: 4005 15-MAY-03 ASE/LL MELP8 MOH/BONE/3M,3F SETI
NUMBER LISTENERS = 8
NUMBER SPEAKERS = 6

COMBINED RESULTS - STANDARD ERRORS ACROSS SPEAKERS AND LISTENERS ******

| MAIN ATTRIBUTE: | PRESENT | | ABSENT | | TOTAL | |
|---|---|---|---|---|---|---|
| | MEAN | S.E. | MEAN | S.E. | MEAN | S.E. |
| VOICING | 95.31 | 1.52 | 80.99 | 2.50 | 88.15 | 1.42 |
| FRICTIONAL | 96.35 | 1.82 | 70.83 | 4.03 | 83.59 | 2.24 |
| NON-FRICTIONAL | 94.27 | 1.86 | 91.15 | 2.04 | 92.71 | 1.17 |
| NASALITY | 96.88 | 0.87 | 75.52 | 3.02 | 86.20 | 1.59 |
| GRAVE | 94.79 | 1.48 | 69.79 | 4.65 | 82.29 | 2.44 |
| ACUTE | 98.96 | 0.73 | 81.25 | 3.20 | 90.10 | 1.66 |
| SUSTENTION | 65.89 | 2.54 | 68.23 | 3.20 | 67.06 | 1.61 |
| VOICED | 71.88 | 4.32 | 57.81 | 6.58 | 64.84 | 3.38 |
| UNVOICED | 59.90 | 3.55 | 78.65 | 3.24 | 69.27 | 2.04 |
| SIBILATION | 79.69 | 2.49 | 75.26 | 2.99 | 77.47 | 1.99 |
| VOICED | 78.65 | 3.41 | 85.42 | 2.21 | 82.03 | 1.89 |
| UNVOICED | 80.73 | 3.27 | 65.10 | 5.41 | 72.92 | 3.08 |
| GRAVENESS | 61.46 | 2.88 | 57.29 | 3.11 | 59.38 | 2.38 |
| VOICED | 78.65 | 2.88 | 79.69 | 2.54 | 79.17 | 2.02 |
| UNVOICED | 44.27 | 4.48 | 34.90 | 5.41 | 39.58 | 3.69 |
| COMPACTNESS | 70.05 | 2.85 | 77.86 | 3.03 | 73.96 | 2.41 |
| VOICED | 85.42 | 2.77 | 79.69 | 3.77 | 82.55 | 2.29 |
| UNVOICED | 54.69 | 4.19 | 76.04 | 4.07 | 65.36 | 3.46 |
| TOTALS | 78.21 | 1.03 | 72.53 | 1.36 | 75.37 | 1.02 |

TOTAL DRT SCORE:
**********
* MEAN = 75.37 *
*
* S.E. = 1.02 *
**********

| | MO | MA | MB | FO | FB | FC |
|---|---|---|---|---|---|---|
| MEAN | 77.99 | 83.20 | 77.86 | 66.28 | 69.14 | 77.73 |
| S.E. | 1.14 | 0.95 | 0.98 | 2.09 | 1.61 | 1.66 |

# Quality - Pair Comparison (A/B) Task Instructions

ARCON Corporation
J.D. Tardelli - jdtardelli@arcon.com

- **In this test we are evaluating systems that might be used for military voice communications under a variety of conditions. You are going to hear a number of samples of speech played to you through your headset. Each sample will consist of a sentence that was produced with two different communication systems. The first is identified as A and the second is identified as B.**
- **Please listen to both A and B and then decide which of the two you prefer. Preference is strictly your decision. Some samples will include background noise. Consider the whole signal, background and speech. Your decision should be based on your comparison of the of the quality of the whole signals.**
- **Some of the A/B pairs will seem clearly different and your decisions will be effortless. Others may be more difficult. ALWAYS MAKE A DECISION BETWEEN THE TWO. "I DON'T KNOW" and "I don't like either one" ARE NOT OPTIONS. Make your decisions independently. You should always compare the two current sentences, and not use any other presentation.**
- **This is not a test of you in any way; it is an evaluation of the systems. There is no right or wrong. Do not discuss how you are making your ratings during breaks or stretching periods.**
- **To indicate your choice, use the button box at your test station. After listening to the two sentences, all lights on the box will flash. At that time, please press the appropriate single button that represents your opinion of the communication quality of the sample just heard. Use the leftmost button if you preferred the first sample (or A). Use the button on the far right if you preferred the second speech sample (or B). The corresponding light will be activated when a choice has been indicated. Once the button has been pushed, you will not be allowed to change your mind, so please respond carefully.**
- **If you attempt to indicate your opinion before the completion of the sentence sequence, when all the lights flash, your choice will not register and no light will be activated. The test will not proceed until an opinion has been indicated for all participants, therefore, be sure that you have a light lit indicating your opinion.**
- **After you have given your opinion there will be a short pause before the next sample begins.**
- **Thank you for participating in this research. Feel free to ask any procedural questions at this time.**
- **For practice, you will hear several samples and give an opinion on each. Then there will be a break to make sure that everything is clear before we start the test. Do not hesitate to ask questions if you have any.**